Studies in Educational Evaluation 82 (2024) 101367

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

How schools affect equity in education: Teaching factors and extended day programs associated with average achievement and socioeconomic achievement gaps

Max Nachbauer

Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

ARTICLE INFO

Keywords:
Teaching quality
All-day school
After-school program
Homework supervision
Tutoring
Educational inequality

ABSTRACT

The current study evaluates effects of teaching factors and extended day programs regarding both average achievement and socioeconomic achievement gaps. Analyses are based on data from the National Educational Panel Study in Germany. The sample comprises 1523 students in 120 classes in 71 secondary schools. The investigation period covers fifth grade to seventh grade. The dependent variable is mathematics achievement. Data are analyzed using intercept-and-slope-as-outcome models. Analyzes on teaching factors reveal that classroom management and cognitive activation increase average achievement. Forming heterogenous groups within the classroom reduces socioeconomic achievement gaps. Analyzes on extended day programs reveal that tutoring reduces socioeconomic achievement gaps. Homework supervision reduces average achievement and increases socioeconomic achievement gaps. Enrichment programs increase socioeconomic achievement gaps. Implications for research, policy and practice are drawn.

1. Introduction

Evaluations of educational effectiveness can be based on two dimensions: the quality and the equity dimension. The *quality dimension* refers to the average level of learning outcomes, while the *equity dimension* refers to differences in learning outcomes between students (Creemers & Kyriakides, 2008; Nachbauer & Kyriakides, 2020). In terms of equity, the goal is that all students learn successfully regardless of their family background (Charalambous, Kyriakides, & Creemers, 2018). However, there is a growing societal awareness that this ideal often is not realized. For instance, numerous studies reveal socioeconomic achievement gaps meaning that students from families with high socioeconomic status (SES) show better academic achievement than students from families with low SES (Liu, Peng, Zhao, & Luo, 2022). Against this backdrop, international institutions like the European Commission (European Commission/EACEA/Eurydice, 2020) or the United Nations (2015) consider equity in education a major political goal.

In order to promote equity in education, a deeper understanding of interventions that effectively reduce socioeconomic achievement gaps is needed. Theoretical models of educational effectiveness generally emphasize the importance of the classroom and the school level (Scheerens, 2013). Regarding the quality dimension, a broad research base demonstrates that classes and schools differ in average achievement (Luyten, 2003; Scheerens & Bosker, 1997). Regarding the equity dimension, the picture is less clear, but there is some evidence that the degree of socioeconomic achievement gaps varies between classes and schools (Gustafsson, Nilsen, & Yang Hansen, 2018; Kyriakides, Creemers, & Charalambous, 2018, 2019b). This suggests that factors at the classroom and the school level affect socioeconomic achievement gaps. In line with this, Kyriakides et al. (2019a) report that a school improvement approach that focuses on school policies for teaching and the school learning environment indeed reduces socioeconomic achievement gaps. Thus, school-based interventions seem generally suitable for promoting equity.

However, it is currently not sufficiently clarified which teaching/school factors play a role for the equity dimension. Most educational effectiveness studies so far focus solely on the quality dimension while only few studies also consider the equity dimension (Atlay, Tieben, Fauth, & Hillmert, 2019b; Nachbauer & Kyriakides, 2020). Therefore, much is known about teaching/school factors affecting academic achievement, but little is known about teaching/school factors affecting achievement gaps (Gustafsson et al., 2018). The current study addresses this topic by investigating socioeconomic achievement gaps in German

E-mail address: max.nachbauer@fu-berlin.de.

https://doi.org/10.1016/j.stueduc.2024.101367
Received 24 November 2023; Received in revised form 17 April 2024; Accepted 23 April 2024
Available online 4 May 2024

secondary schools. Specifically, equalizing effects of teaching factors and extended day programs on gaps in mathematics achievement are examined.

2. Background on socioeconomic achievement gaps

Equity in education can be conceptualized in different ways (Charalambous et al., 2018; Nachbauer & Kyriakides, 2020). From a *strict egalitarian view*, a reduction of differences in learning outcomes between students is generally viewed as desirable (Helmke, 1988). However, it seems questionable whether this strict view truly reflects fairness as differences in learning outcomes may have legitimate causes (e.g., differences in students' effort). In light of this, a *moderate egalitarian view* has evolved. From this view (Charalambous et al., 2018), specifically those differences in learning outcomes that are related to students' background characteristics (e.g., social background, migration background, gender) should be kept to a minimum. The rationale behind this view is that effects of background characteristics are considered problematic because students are unable to influence these factors. The moderate egalitarian view is widely accepted in the literature (e.g., Atlay et al., 2019b; Fischer, Theis, & Züchner, 2014; Gustafsson et al., 2018) and is also used in the current study. Specifically, the study focuses on achievement gaps related to SES.

Essential indicators of family SES are parents' educational attainment, occupational status, and income. A broad research base demonstrates that family SES affects students' academic achievement. In a meta-analysis Liu et al. (2022) found a mean correlation between SES and academic achievement of $r = 0.28$. Thus, SES constitutes a major influence on educational success. This stands in stark contrast to the notion of equity in education. Therefore, it seems necessary to find ways to reduce socioeconomic achievement gaps. The theory-driven identification of school-based interventions builds on reflections about the causes of socioeconomic achievement gaps. If low-SES students are to be supported, the reasons for their low academic achievement need to be understood.

The emergence of socioeconomic achievement gaps is a scientific field that touches several disciplines, including educational research, educational psychology, sociology of education, and economy of education. Recently, Nachbauer (2023) integrated these different perspectives in a theoretical model. The model of mediating mechanisms between social background and learning outcomes is presented in Fig. 1. It describes a multi-step mediation process: social background is associated with family resources, which affect learning environments, which in turn affect learning processes.

The conceptualization of *family resources* is based on Brophy and Good (1986) forms of capital but focuses more strongly on factors related to learning. Family resources comprise economic resources (e.g., expenses on children, place of residence), personal resources (e.g., parents' skills, parents' educational aspirations) and social resources (e.g., parent to child ratio, family social network). These resources form the basis for children's education. Therefore, they are the starting point for achievement gaps: high-SES families dispose of considerably more resources than low-SES families (for an overview see Nachbauer, 2023).

Due to differences in family resources, children with various social backgrounds experience different *learning environments*. High-SES children grow up in more stimulating home learning environments than low-SES children, meaning that high-SES parents interact with their children in a more beneficial way than low-SES parents (Cooper, Lindsay, & Nye, 2000; Hoff, Laursen, & Tardif, 2002). Further, high-SES children and low-SES children typically differ regarding the attendance of educational facilities, including early childhood facilities, schools, school-related facilities (e.g., private tutoring) and recreational facilities (e.g., music school). This is mainly due to differences in parents' educational choices (Boudon, 1974) meaning that high-SES parents evaluate the costs and returns of educational participation differently than low-SES parents. Notably, socioeconomic differences in educational participation vary between countries indicating that system level moderators play a role, too (OECD, 2016; Zwier, Geven, & van de Werfhorst, 2021).

Differences in learning environments result in differences in *learning processes*. Thus, children with various social backgrounds have different opportunities to develop certain learning prerequisites (e.g., cognitive characteristics, personality characteristics). For instance, high-SES children have more prior knowledge, better metacognitive skills, and better self-regulation skills than low-SES children (Garcia, 2015; Karlen, Maag Merki, & Ramseier, 2014; Miech, Essex, & Goldsmith, 2001). Therefore, it is much easier for high-SES children to meet the challenges of school learning.

Concerning interventions for the reduction of achievement gaps, several conclusions can be drawn from the theoretical model. First, it should be emphasized that teaching/school factors might work as mediating mechanisms between SES and academic achievement. Thus, if schools aim at reducing achievement gaps, they first need to understand to which extent they themselves contribute to the emergence of these achievement gaps. Second, children within the same classroom differ in terms of their learning environments outside the school. Particularly, parent-child-interactions vary between high-SES families and low-SES families. These family factors might interact with teaching/

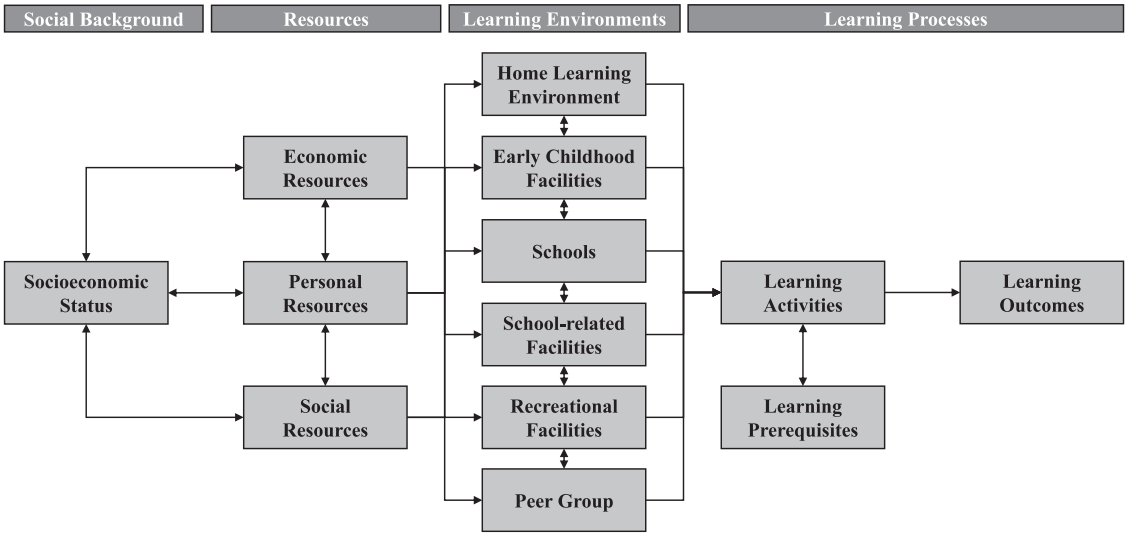

Fig. 1. Theoretical model of mediating mechanisms between social background and learning outcomes (Nachbauer, 2023, p. 54).

school factors. Third, children within the same classroom differ in terms of their learning prerequisites, with high-SES children having more favorable learning prerequisites than low-SES children. Again, these learning prerequisites might interact with teaching/school factors.

3. Background on teaching/school factors associated with socioeconomic achievement gaps

In the following, mechanisms underlying interactions between teaching/school factors and social background are described (Section 3.1) and the compensatory potential of teaching factors (Section 3.2) and extended day programs (Section 3.3) is elaborated on.

3.1. Mechanisms underlying the effects of teaching/school factors on socioeconomic achievement gaps

Basically, there are two ways to attain a reduction of socioeconomic achievement gaps: either the achievement gains of low-SES students are increased, or the achievement gains of high-SES students are reduced. While in both cases the strength of the socioeconomic achievement gap is reduced, obviously only the first case is pedagogically sensible (Baumert & Schümer, 2001). Thus, equity in education is not a stand-alone goal, but needs to be in accordance with the quality dimension. This means that schools that reduce achievement gaps by increasing low-SES students' achievement are desirable while the same is not true for schools that reduce achievement gaps by decreasing high-SES students' achievement (Nuttall, Goldstein, Prosser, & Rasbash, 1989). Increasing low-SES students' achievement is especially worthwhile because it makes equity and quality complementary goals. Indeed, there is some evidence that schools that are effective at reducing socioeconomic achievement gaps are also effective at increasing average achievement (Kyriakides et al., 2018, 2019b). Therefore, it seems possible to promote the equity dimension and thereby also the quality dimension (Kyriakides et al., 2019b).

For theoretical reflections on factors that might affect socioeconomic achievement gaps, the complex interrelationship between school and family needs to be understood. Three patterns of how teaching/school factors and family background might interact can be distinguished (Nachbauer, 2023). These patterns are outlined in Fig. 2. For illustration purposes, the issue is explained using the example of tutoring, academic achievement, and SES.

A teaching/school factor has an *effect on inequality in learning outcomes* when different manifestations of the teaching/school factor result in different degrees of inequality in learning outcomes. This would be the case for tutoring when socioeconomic achievement gaps are lower in schools that offer a tutoring program than in schools that do not offer a tutoring program. The upper part of Fig. 2 provides a graphical illustration of an effect on inequality. Statistically speaking, this is a moderation, more specifically a cross-level interaction. This means that the teaching/school factor moderates the relationship between family background and learning outcome.

Such an effect on inequality in learning outcomes may be due to two mechanisms: differentially effective learning opportunities or different learning opportunities. A *differentially effective learning opportunity* is given when the same learning opportunity produces different learning outcomes for students with various family backgrounds. In the example, the equalizing effect of tutoring might emerge because the effect of tutoring on academic achievement is stronger for the participating low-SES students than for the participating high-SES students. Statistically, this is a moderation, too. However, the interaction is now located at the student level and family background moderates the relationship between teaching/school factor and learning outcome (see middle part of Fig. 2).

Alternatively, an effect on inequality in learning outcomes might be due to a second mechanism: *different learning opportunities*. These are given when students with various family backgrounds experience systematically different learning opportunities in schools (for empirical results see Atlay, Tieben, Fauth, & Hillmert, 2019a; Fauth, Atlay, Dumont, & Decristan, 2021; Opdenakker & Van Damme, 2001; Rjosk, Richter, Hochweber, Lüdtke, Klieme, & Stanat, 2014). This means that there is a relationship between family background and the manifestation of the teaching/school factor at the student level. This can be illustrated with the example of tutoring. The equalizing effect of tutoring might emerge because low-SES students more often participate in the school-based tutoring program than high-SES students. In terms of statistics, this is a mediation. Family background affects the teaching/-school factor which in turn affects the learning outcome.

3.2. Effects of teaching factors on achievement and achievement gaps

The previously described mechanisms can be used for the theory-driven identification of teaching/school factors that might affect socioeconomic achievement gaps. This section focuses on teaching factors while the next section focuses on extended day programs.

The assumption that teaching quality is associated with socioeconomic achievement gaps can be frequently found in the literature (Atlay et al., 2019b; Nilsen, Scherer, Gustafson, Teig, & Karstein, 2020). However, teaching quality is a multidimensional construct, and its various aspects might affect achievement gaps differently (Atlay et al., 2019b). Within the German context, teaching quality is often described using the three basic dimensions classroom management, support, and cognitive activation (Praetorius, Klieme, Herbert, & Pinger, 2018). *Classroom management* refers to teacher behaviors that aim at maximizing students' active learning time in class. This comprises classroom organization as well as management of disruptive behavior of students. Classroom management is known to enhance academic achievement (Marzano, Marzano, & Pickering, 2003). Further, the aspect of behavioral management might play a role for socioeconomic achievement gaps. Classroom management might be differentially effective as low-SES students have lower self-regulation than high-SES students (Miech et al., 2001) and therefore are in greater need of external regulation. Some empirical results are in line with this assumption (Brophy & Good, 1986; Palardy, 2008), while other studies do not confirm equalizing effects of classroom management (Atlay et al., 2019b).

The dimension *support* refers to the quality of social interactions between teacher and students as well as between students and students (Praetorious et al., 2018). Support by teachers can be further differentiated into the subdimensions academic support, social-emotional support, and autonomy support (Tao, Meng, Gao, & Yang, 2022). The current study focuses on teachers' academic support. Relevant teacher behaviors include assessment of learning processes, promotion of a positive error climate, and providing cues, prompts and feedback. It seems plausible that teachers' academic support affects socioeconomic achievement gaps (Atlay et al., 2019b). Teachers' academic support might be differentially effective as low-SES students have deficits in prior knowledge and metacognitive skills and receive inadequate support from their parents (Cooper et al., 2000; Garcia, 2015; Karlen et al., 2014). Moreover, the mechanism of different learning opportunities is thinkable meaning that teachers might provide more academic support for low-SES students than for high-SES students (Atlay et al., 2019a). However, empirical results are quite inconsistent. Early studies reported that proactive teacher support and praise are especially relevant for low-SES students (Brophy & Good, 1986). In contrast, a more recent study found that a supportive climate is more beneficial for high-SES students (Atlay et al., 2019b).

Cognitive activation refers to instructional practices that engage students in higher level thinking processes (Lipowsky et al., 2009). This includes aspects like discussion of relationships between concepts, stimulation of cognitive conflicts, setting challenging tasks and comparing different solution strategies. Several studies provide evidence for positive effects of cognitive activation on academic achievement, especially in the field of mathematics learning (Baumert et al.,

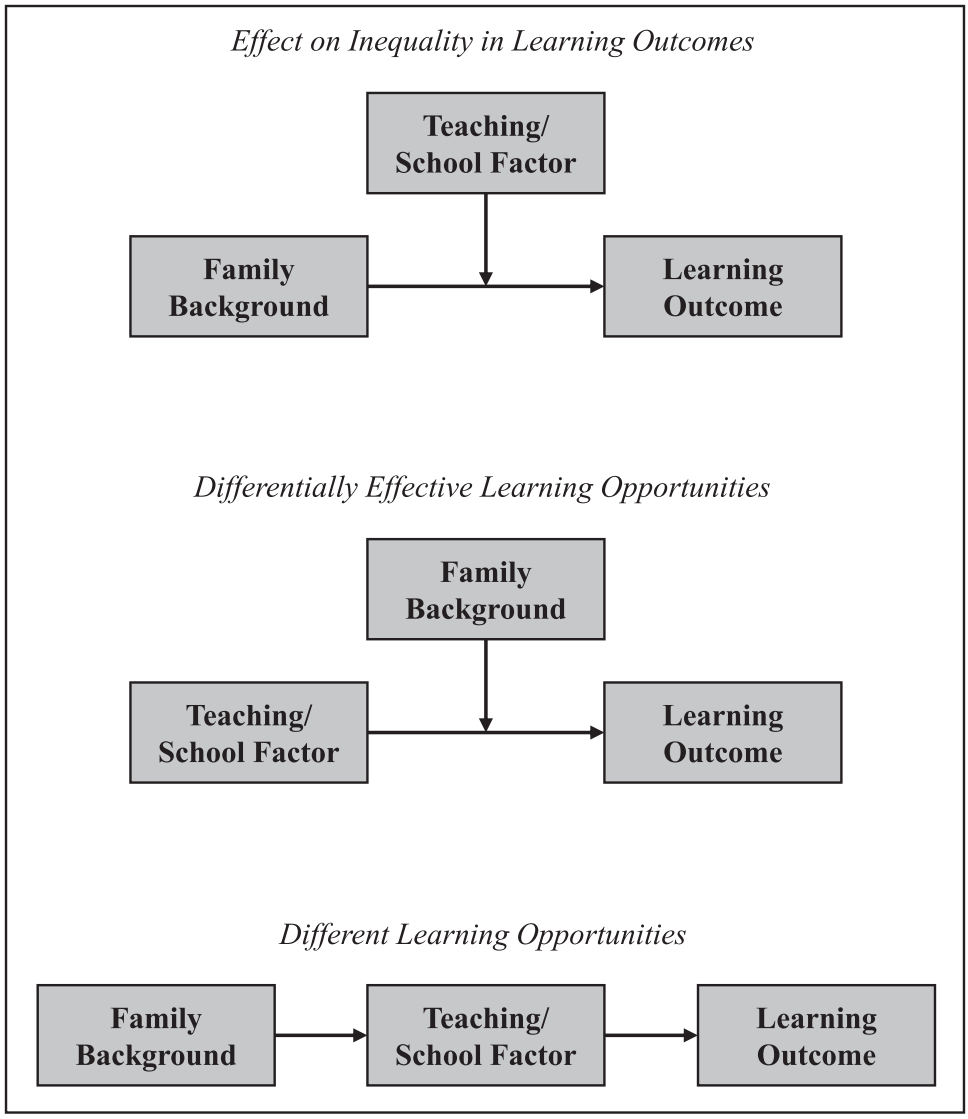

Fig. 2. Possible relationships between teaching/school factors, family background and learning outcomes (Nachbauer, 2023, p. 112).

2010; König et al., 2021; Lipowsky et al., 2009). Noteworthy, a recent cross-country study reported that in most countries a curvilinear association between cognitive activation and achievement exists (Caro, Lenkeit, & Kyriakides, 2016). This suggests that a moderate degree of cognitive activation is most beneficial for learning. Regarding the equity dimension, differential effects of cognitive activation are thinkable as these instructional practices might put high requirements on students' prior knowledge and metacognitive skills. Indeed, some studies find that effects of cognitive activation are stronger for high-SES students than for low-SES students (Atlay et al., 2019b; Caro et al., 2016).

While the three basic dimensions are of central importance for teaching quality, there is consensus that further teaching factors are relevant for learning (Praetorius et al., 2018). One teaching factor that should additionally be considered is the *grouping* of students within the classroom. Grouping refers to situations when students learn together in the form of partner work or group work. In terms of equity in education, it seems plausible that the specific type of group composition plays a role. With *homogenous groups*, high-achieving students have more stimulating interaction partners than low-achieving students. With *heterogenous groups*, interaction partners are distributed equally. Still, it can be expected that heterogenous groups are differentially effective. For low-SES students, interactions with other students might be especially beneficial as peers provide additional academic support. This assumption is confirmed by several studies: effects of heterogenous groups are stronger for disadvantaged students than for privileged students (Lou et al., 1996; Rohrbeck et al., 2003).

3.3. *Effects of extended day programs on achievement and achievement gaps*

While regular class is at the core of schooling, further learning opportunities exist in schools. In this study, the term extended day program refers to educational offers that schools provide in the afternoon, subsequent to regular class.[1] In the discussion on extended education, different concepts are used (e.g., all-day schools, out-of-school time, after-school program), which are dependent on national contexts and their educational policies. The following considerations focus on the German context.

[1] It should be noted that in the international literature the term extended day program often refers to programs that increase the allocated class time (e.g., Meyer & Van Klaveren, 2013). However, this is not a necessary constituent in German all-day schools. In many German all-day schools, the extended day program comprises only offers that supplement regular class (e.g., homework supervision, sports programs, arts programs).

In Germany, the common way of organizing schools for a long time has been half-day schools. A *half-day school* provides only regular class and is closed in the afternoon. However, in the last 20 years, a strong expansion of all-day schools took place. An *all-day school* is defined as a school that provides an extended day program at least three days per week. Three types of all-day schools are distinguished. In *open all-day schools,* students can freely decide to attend the extended day program or not. In *partially bound all-day schools,* a part of the students is obliged to attend the extended day program. In *fully bound all-day schools,* all students are obliged to attend the extended day program.

The expansion of all-day schools in Germany was driven by political intentions of widening childcare services, promoting students' learning outcomes and reducing socioeconomic achievement gaps (Fischer et al., 2014). Differential effects of all-day schools are expected primarily because low-SES students experience less stimulating home learning environments than high-SES students. Yet, most empirical evaluations so far yield sobering results: compared to half-day schools, all-day schools attain a similar average level of academic achievement and a similar strength of socioeconomic achievement gaps (Linberg, Struck, & Bäumer, 2018; Strietholt, Manitius, Berkemeyer, & Bos, 2015). Only one study indicates that types of all-day schools differ regarding the equity dimension: socioeconomic achievement gaps are lower in fully bound all-day schools than in open or partially bound all-day schools (Fischer et al., 2014). This suggests that differences in the participation in extended day programs between high-SES students and low-SES students might play a role.

Conceptually, types of all-day schools are considered a structural factor meaning that they determine basic conditions for afternoon learning. In contrast, *types of extended day programs* are process factors in that they describe the actual learning activities that occur. The current study focuses on extended day programs that emphasize academic learning. Many schools offer *homework supervisions* where school staff looks after students while they are working on their homework. Despite their high prevalence, several studies find no effects of homework supervisions on academic achievement (Cosden, Morrison, Gutierrez, & Brown, 2004; Steinmann, Strietholt, & Caro, 2019). A conceptually related extended day program is tutoring. *Tutoring* refers to one-on-one or small-group remedial instruction that supplements regular classroom instruction (Nickow, Oreopoulos, & Quan, 2020). Effects of tutoring are well-researched. A recent meta-analysis confirms positive effects of school-based tutoring on academic achievement (Nickow et al., 2020).

In terms of the equity dimension, it seems plausible that homework supervision and tutoring may reduce socioeconomic achievement gaps, as both extended day programs provide academic support. These extended day programs might have differential effects for low-SES students and high-SES students. This assumption is based on the same arguments that were already mentioned in the section on teacher support in class: low-SES students have deficits in prior knowledge and metacognitive skills and receive inadequate support from their parents. Indeed, there is evidence that tutoring has stronger effects for disadvantaged students (Fryer & Howard-Noveck, 2020). Further, the mechanism of different learning opportunities might be relevant. As low-SES students face more learning problems than high-SES students, they more often participate in homework supervision and tutoring (Linberg, Struck, & Bäumer, 2015; Nachbauer, 2023).

While tutoring focuses on low-achieving students, there are also extended day programs that focus on high-achieving students. In *enrichment programs* high-achieving students learn a subject on an advanced level by receiving additional content with more breadth and/or more depth (Kim, 2016). Enrichment activities prove to foster the academic achievement of the participating students (Kim, 2016). However, given that low-SES students are usually underrepresented in enrichment programs (different learning opportunities), it might be that enrichment programs increase socioeconomic achievement gaps.

4. Research questions

Socioeconomic achievement gaps remain a major challenge for educational policy and practice. Thus, there is a strong need for school-based interventions to reduce socioeconomic achievement gaps. However, the empirical knowledge base on this topic is still insufficient. The number of existing studies is limited and in several domains the findings are quite inconsistent. Therefore, this study evaluates the effectiveness of different teaching factors and types of extended day programs regarding both the quality and the equity dimension. Four research questions are examined:

- RQ1) What are the effects of teaching factors on the average level of academic achievement?
- RQ2) What are the effects of teaching factors on socioeconomic achievement gaps?
- RQ3) What are the effects of extended day programs on the average level of academic achievement?
- RQ4) What are the effects of extended day programs on socioeconomic achievement gaps?

Based on the previous theoretical considerations and the reviewed empirical findings, it is hypothesized that classroom management and cognitive activation both increase average achievement. Further, classroom management, academic support and heterogenous groups should reduce socioeconomic achievement gaps while cognitive activation should increase socioeconomic achievement gaps. Regarding extended day programs, it is hypothesized that fully bound all-day schools, tutoring and homework supervision should reduce socioeconomic achievement gaps while enrichment programs should increase socioeconomic achievement gaps.

5. Method

5.1. Data

Empirical analyses are based on data from the National Educational Panel Study (NEPS). The NEPS is a large-scale multi-cohort research project in Germany (Blossfeld & Rossbach, 2019). It covers six starting cohorts that are examined longitudinally (1. newborns, 2. kindergarten, 3. grade five, 4. grade nine, 5. university students, 6. adults). All samples are representative for the whole of Germany.

The current study analyzes starting cohort 3 (NEPS Network, 2023). This starting cohort begins in grade five, in which students usually enter secondary schools in Germany. The sample was drawn using a stratified cluster sampling procedure. The full sample including all types of schools comprises 6112 students. Analyses are restricted to students within academic track secondary schools (Gymnasium). Academic track secondary schools were favored over other types of secondary schools because student heterogeneity in terms of SES is most pronounced in this type of secondary school (Baumert & Schümer, 2001; Nachbauer, 2023). The sample of academic track secondary schools comprises 2415 students.

Due to the focus of the study, a part of the original sample was excluded. Students with missing values on the SES variable were excluded.[2] This was necessary because the used method for handling

[2] The vast majority of excluded students are students with missing values on the SES variable. The large proportion of missing values on this variable is caused by the methodological approach of the NEPS. Within the NEPS, families can freely choose in which measurements they want to participate and in which measurements not (student achievement tests, student questionnaires, parent telephone interviews). A considerable number of parents permits student achievement tests and students questionnaires but denies telephone interviews. Thus, these cases have missing values on the SES variable.

missing data (full information maximum likelihood estimation) could not be applied to the SES variable (due to random slope modeling). Further, classes with less than five students participating in the NEPS were excluded. This approach was chosen because a sufficient number of students per class is critical for modeling within-class achievement gaps (Nachbauer & Kyriakides, 2020). The final analytical sample comprises 1523 students in 120 classes in 71 schools. The analytical sample contains 63% of the original sample of academic track secondary schools and does not differ considerably from the original sample (for comparative analyses see Nachbauer, 2023).

5.2. Variables

In Germany, students generally enter academic track secondary schools in grade 5. Data were collected at three waves: at the outset of grade 5 (end of 2010), at the outset of grade 6 (end of 2011) and at the outset of grade 7 (end of 2012). Data were collected using standardized achievement tests, telephone interviews with parents, teacher questionnaires and principal questionnaires. All items of the teacher and principal questionnaires can be found in the Appendix, section A.

5.2.1. Student level variables

The dependent variable is students' mathematics achievement in grade 7. Within the NEPS, mathematics achievement is measured with a standardized achievement test that covers several content areas (e.g., quantity, change and relationship, space and shape) and several cognitive processes (e.g., mathematical argumentation, mathematical modeling, solving mathematical problems). Mathematics achievement is measured at the beginning of grade 5 and at the beginning of grade 7. Therefore, the variable *mathematics prior achievement* refers to grade 5 while the variable *mathematics final achievement* refers to grade 7. Both variables are weighted likelihood estimates calculated with item response theory models. These variables are provided by the NEPS.

Family *SES* is operationalized using parents' educational attainments, gathered with telephone interviews with parents. The measure is based on the CASMIN classification (Brauns, Scherer, & Steinmann, 2003), which covers both school-leaving qualifications and vocational qualifications. Educational attainments according to the CASMIN classification were converted to equivalent years of education. Equivalent years of education refer to the number of years that are usually needed to accomplish a specific educational attainment. For instance, the lowest school-leaving qualification (Hauptschulabschluss) in combination with no vocational qualification equates to 9 years of education. The highest school-leaving qualification (Abitur) in combination with a master's degree equates to 18 years of education. The variable on CASMIN years of education is provided by the NEPS. It was chosen for this study because it enables measuring parental education on a metric scale. In order to have one SES value per student, the highest educational attainment in the family is used (the same approach is applied by Atlay et al., 2019b).

Further, students' *gender* (0 = boy, 1 = girl) and *migration background* (0 = no, 1 = yes) are considered as control variables. A migration background is given when at least one parent is born outside of Germany. This information is gathered with telephone interviews with parents.

5.2.2. Teaching factors

All variables on teaching factors are gathered with questionnaires for mathematics teachers. *Classroom management* refers to the proportion of time that the teacher on average spends with maintaining order and dealing with interruptions. The variable was inverted ($x' = 1 - x$), thus a higher value refers to a more efficient classroom management. Two variables are related to teachers' academic support. *Tasks with academic support* refer to the proportion of time that the teacher on average lets students work on tasks while he/she provides assistance. In order to have a contrast, a second variable is used. *Tasks without academic support* refer to the proportion of time that the teacher on average lets students work on tasks while he/she does not provide assistance. *Cognitive activation* is measured with a scale comprising four items. The items refer to the types of assignments that the teacher gives to students (e.g., assignments that require time to think, assignments that involve selection of the right approach). The factorial structure and internal consistency of this scale was established in a previous study (Nachbauer, 2023). Finally, two variables are related to how the teacher groups students within the classroom. *Homogenous groups* refer to forming groups of students with similar achievement levels while *heterogenous groups* refer to forming groups of students with different achievement levels.

5.2.3. Extended day programs

Most of the variables on extended day programs are gathered with questionnaires for principals. *Type of school* is a structural factor related to extended day programs. It is a categorical variable comprising five different types of school. Five dummy variables were created: *half-day school*, *half-day school with limited extended day program* (extended day program 1–2 days per week), *open all-day school*, *partly bound all-day school,* and *fully bound all-day school*. All further variables on extended day programs are process factors. *Homework supervision* refers to how often the school offers a homework supervision. Accordingly, *enrichment programs* refer to how often the school offers enrichment teaching for high-achieving students. In addition to the variables from the questionnaire for principals, there is one variable coming from the questionnaire for mathematics teachers. *Tutoring* refers to the amount of additional remedial mathematics lessons per week.

5.2.4. Class composition

In order to control for effects of student composition, the means of mathematics prior achievement, SES, migration background and gender at the class level were calculated. However, controlling for student composition did not affect the results on teaching factors and extended day programs. Therefore, in the final models only the variable *mean SES* is included. This variable is relevant for the chosen centering approach (see next section).

5.2.5. Centering variables

All continuous variables were centered at the grand-mean to facilitate interpretation of coefficients (Enders & Tofighi, 2007). Categorical variables (gender, migration background, type of school) were not centered. As this study examines within-class achievement gaps, the relationship between SES and mathematics final achievement at the student level is of interest. Estimating the student level relationship is realized by grand-mean centering SES and modeling a compositional effect of mean SES (Enders & Tofighi, 2007). A further explanation of the rationale behind this centering approach can be found in Appendix, section B.

5.3. Analytical strategy

For analysis of data, multilevel modeling is applied. Specifically, an intercept-and-slope-as-outcome model is used (Raudenbush & Bryk, 2002). Within this model, the random intercept is an indicator for the quality dimension, while the random slope of SES is an indicator for the equity dimension. In the student level part of the model, the predictors are students' mathematics prior achievement, SES, migration background and gender. Thus, the random intercept indicates students' *average achievement gains* adjusted for students' background characteristics (value added model, see Sammons, 1999). Regarding the equity dimension, controlling for prior achievement is of central importance, too. This is because schools may differ in prior achievement gaps between students (Nachbauer & Kyriakides, 2020). When students enter school, in some schools there might already be large socioeconomic achievement gaps (e.g., because the school catchment area is socially heterogeneous) while in other schools there might be only small

socioeconomic achievement gaps (e.g., because the school catchment area is socially homogenous). This needs to be taken into account when school effects on equity are to be estimated. One suitable approach is controlling for students' prior achievement (Nachbauer & Kyriakides, 2020). Thus, the random slope indicates *socioeconomic gaps in achievement gains*.

Regarding the upper levels of the model, it should be noted that some independent variables are measured at the class level (variables on teaching factors and tutoring from the mathematics teacher questionnaire) while other independent variables are measured at the school level (variables on extended day programs from the principal questionnaire). However, it was not possible to model both class level and school level, due to the sample structure (120 classes in 71 schools). Therefore, it was decided to use a two-level model that includes the student level and the class level. Variables measured at the school level were disaggregated at the class level.[3]

Testing effects of teaching factors/extended day programs included linear and quadratic effects. Quadratic effects examine whether a factor has a curvilinear relationship (inverted-U) to the learning outcome, meaning that a moderate degree of the factor is optimal for learning (Creemers & Kyriakides, 2008). While for all continuous variables on teaching factors/extended day programs quadratic effects were tested, only significant quadratic effects are reported in the results section, for the sake of a concise presentation.

Statistical analyses were conducted using Mplus 8.2 (Muthén & Muthén, 1998–2018). Given that some variables are not normally distributed a robust maximum likelihood estimator was applied. Missing values were treated with full information maximum likelihood estimation.

6. Results

Descriptive statistics for all variables can be found in the Appendix, section C. It should be noted that the SES variable has a left-skewed distribution. This is due to sample of academic track secondary schools. Therefore, the terms *higher-SES students* and *lower-SES students* are used subsequently. These terms highlight the reference norm of academic track secondary schools.

In an empty random intercept model, the variance of the random intercept is significant ($p < 0.05$). 86.3% of the total variance of mathematics final achievement is located at the student level and 13.7% at the class level. This means that the average level of achievement varies between classes. In the next step, SES is added as a predictor at the student level and modeled with a random slope. The variance of the random slope is significant ($p < 0.05$). This means that the strength of socioeconomic achievement gaps varies between classes. The correlation between random intercept and random slope is negative ($r = -0.354$), albeit not significant ($p > 0.10$). Thus, there is a tendency that a higher average level of achievement comes along with lower socioeconomic achievement gaps. Even though the correlation is not significant, it is clear that the quality dimension and the equity dimension are not conflicting.

In the next step, mathematics prior achievement, migration background and gender are added as predictors at the student level. This model is hereafter referred to as the baseline model. Regarding the effects of the student level predictors, the estimates are quite similar across the different models (see Tables 1 and 2). All effect estimates are unstandardized regression coefficients (because standardized regression coefficients are not available in random slope models). Mathematics prior achievement has a significant positive effect on mathematics final achievement. SES has, on average, a significant positive effect on mathematics final achievement. This means that higher-SES students have a higher relative achievement gain than lower-SES students. Gender has a significant negative effect on mathematics final achievement. This means that boys have a higher relative achievement gain than girls. The effect of migration background is not significant. In all models, the residual variance of mathematics final achievement at the student level (which is the variance that is not explained by the student level predictors) remains significant.

Finally, predictors at the class level are added. In each presented model (Model 1a – Model 2d), one predictor is modeled at the class level. Model fit is evaluated with deviance tests. The deviances of the models with a predictor at the class level are compared with the deviance of the baseline model (only predictors at the student level). The deviance tests are significant for all models (see bottom rows of Tables 1 and 2). This means that all models with a predictor at the class level fit the data significantly better than the baseline model.

6.1. Effects of teaching factors

Results on effects of teaching factors are presented in Table 1. Each column presents the results of an intercept-and-slope-as-outcome model with one teaching factor modeled as a predictor at the class level. Estimates of effects of teaching factors on average achievement gains (quality dimension, RQ1) can be found in the section "Predictors class level: intercept". The effect estimates in the tables are unstandardized regression coefficients. *Classroom management* has a significant positive effect on the random intercept (Model 1a). This means that an efficient classroom management is associated with higher achievement gains. In classes with classroom management 1 standard deviation (SD) above the mean the achievement gains are 0.498 SD higher than in classes with average classroom management. A further insight from Table 1 is that *cognitive activation* has both a significant positive linear and a significant negative quadratic effect on the random intercept (Model 1d). Thus, more cognitive activation generally comes along with higher achievement gains (linear effect). However, beyond an optimal level cognitive activation becomes dysfunctional so that achievement gains are then reduced (quadratic effect). In classes with cognitive activation 1 SD above the mean the achievement gains are 0.253 SD higher than in classes with average cognitive activation. But in classes with cognitive activation 1.5 SD above the mean the achievement gains are only 0.115 SD higher than in classes with average cognitive activation, highlighting the curvilinear relationship. The effects of all other teaching factors on the random intercept are not significant. In most models, the residual variance of mathematics final achievement at the class level (which is the variance of the random intercept that is not explained by the class level predictors) remains significant. Only in the model containing cognitive activation (Model 1d) the residual variance of mathematics final achievement at the class level is not significant any more.

Estimates of effects of teaching factors on socioeconomic gaps in achievement gains (equity dimension, RQ2) can be found in the section "Predictors class level: slope". These coefficients are so called cross-level interactions. *Tasks with academic support* have a marginally significant ($p < 0.10$) negative effect on the random slope (Model 1b). Thus, there is a tendency that more time for tasks with academic support by the teacher comes along with smaller socioeconomic gaps in achievement gains. In classes with tasks with academic support 1 SD above the mean the socioeconomic gaps in achievement gains are 0.100 SD smaller than in classes with average tasks with academic support. Moreover, *heterogenous groups* have a significant negative effect on the random slope (Model 1 f). This means that forming groups of students with different achievement levels is associated with smaller socioeconomic gaps in achievement gains. In classes with heterogenous groups 1 SD above the mean the socioeconomic gaps in achievement gains are 0.128 SD smaller than in classes with average heterogenous groups. The effects of

[3] It was checked whether this procedure biases the estimates of the variables measured at the school level, but this was not the case. For all variables measured at the school level the effect estimates in two-level models that include student level and class level are comparable to effect estimates in two-level models that include student level and school level.

Table 1
Intercept-and-slope-as-outcome models for effects of teaching factors on mathematics final achievement (unstandardized regression coefficients with standard errors in brackets).

	Model 1a	Model 1b	Model 1c	Model 1d	Model 1e	Model 1f
Intercept	1.626 * (0.032)	1.623 * (0.032)	1.633 * (0.033)	1.701 * (0.032)	1.624 * (0.033)	1.640 * (0.033)
Predictors student level						
SES (mean slope)	0.040 * (0.011)	0.040 * (0.011)	0.039 * (0.011)	0.045 * (0.014)	0.040 * (0.011)	0.038 * (0.011)
Mathematics prior achievement	0.538 * (0.026)	0.539 * (0.026)	0.537 * (0.027)	0.539 * (0.026)	0.537 * (0.026)	0.537 * (0.026)
Migration background	-0.013 (0.064)	-0.010 (0.064)	-0.013 (0.064)	-0.002 (0.064)	-0.013 (0.064)	-0.016 (0.064)
Gender	-0.232 * (0.047)	-0.227 * (0.048)	-0.233 * (0.046)	-0.244 * (0.048)	-0.232 * (0.047)	-0.232 * (0.047)
Predictors class level: intercept						
Classroom management	0.033 * (0.013)					
Tasks with academic support		0.003 (0.007)				
Tasks without academic support			0.000 (0.005)			
Cognitive activation				0.134 * (0.046)		
Cognitive activation^2				-0.078 * (0.030)		
Homogeneous groups					0.046 (0.039)	
Heterogenous groups						0.028 (0.058)
Mean SES	0.049 (0.041)	0.042 (0.040)	0.050 (0.040)	0.006 (0.029)	0.046 (0.039)	0.044 (0.042)
Predictors class level: slope						
Classroom management	-0.002 (0.005)					
Tasks with academic support		-0.004$^+$ (0.002)				
Tasks without academic support			0.002 (0.001)			
Cognitive activation				-0.014 (0.012)		
Cognitive activation^2				-0.002 (0.007)		
Homogenous groups					-0.004 (0.019)	
Heterogenous groups						-0.032 * (0.003)
Variance components						
Residual variance mathematics final achievement student level	0.621 * (0.028)	0.621 * (0.028)	0.621 * (0.028)	0.620 * (0.028)	0.620 * (0.028)	0.616 * (0.028)
Residual variance mathematics final achievement class level	0.037 * (0.019)	0.048 * (0.023)	0.048 * (0.022)	0.019 (0.012)	0.045 * (0.022)	0.050 * (0.022)
Residual variance slope	0.001 (0.002)	0.000 (0.002)	0.001 (0.002)	0.000 (0.001)	0.001 (0.002)	0.000 (0.003)
Covariance intercept slope	-0.002 (0.003)	-0.002 (0.003)	-0.003 (0.003)	0.000 (0.003)	-0.002 (0.003)	-0.001 (0.004)
Deviance test	97.685 *	113.611 *	124.340 *	132.748 *	72.125 *	77.441 *

Notes. $^+p < 0.10$, *$p < 0.05$.

all other teaching factors on the random slope are not significant. In all models, the residual variance of the random slope (which is the variance of the random slope that is not explained by the class level predictors) is not significant.

6.2. Effects of extended day programs

Results on effects of extended day programs are presented in Table 2. Estimates of effects of extended day programs on average achievement gains (quality dimension, RQ3) can be found in the section "Predictors class level: intercept". *Homework supervision* has a significant negative effect on the random intercept (Model 2c). This means that a higher frequency of homework supervision is associated with lower average achievement gains. In classes with homework supervision 1 SD above the mean the achievement gains are 0.373 SD lower than in classes with average homework supervision. In a further analysis (not in the table) the correlation between homework supervision and student composition in terms of prior achievement is checked. This correlation is not significant ($p > 0.10$). Thus, there are no indications of reverse causality (schools with many students with low prior achievement do not offer homework supervision more frequently). *Enrichment programs* have a marginally significant ($p < 0.10$) negative effect on the random intercept (Model 2d). A higher frequency of enrichment programs tends to come along with lower average achievement gains. In classes with enrichment programs 1 SD above the mean the achievement gains are 0.282 SD lower than in classes with average enrichment programs. The effects of type of school and tutoring on the random intercept are not significant. In all models, the residual variance of mathematics final achievement at the class level remains significant.

Estimates of effects of extended day programs on socioeconomic gaps in achievement gains (equity dimension, RQ4) can be found in the section "Predictors class level: slope". Some *types of schools* have marginally significant ($p < 0.10$) positive effects on the random slope (Model 2a). Thus, there is a tendency that socioeconomic gaps in achievement gains are larger in *half-day schools with limited extended day program* (0.404 SD larger), *open all-day schools* (0.400 SD larger) and *partly bound all-day schools* (0.386 SD larger) than in *half-day schools*. In contrast, socioeconomic gaps in achievement gains are similar in *fully bound all-day schools* and *half-day schools*. Further, different types of extended day programs have specific effects. *Tutoring* has a significant negative effect on the random slope (Model 2b) meaning that a higher frequency of tutoring is associated with smaller socioeconomic gaps in achievement gains. In classes with tutoring 1 SD above the mean the socioeconomic gaps in achievement gains are 0.117 SD smaller than in classes with average tutoring. Finally, *homework supervision* and *enrichment programs* both have significant positive effects on the random slope (Model 2c and 2d). Thus, a higher frequency of homework supervision and enrichment programs, respectively, come along with larger socioeconomic gaps in achievement gains. In classes with homework supervision (enrichment programs) 1 SD above the mean the socioeconomic gaps in achievement gains are 0.140 (0.097) SD larger than in classes with average homework supervision (enrichment programs). In all models, the residual variance of the random slope is not significant.

7. Discussion

The current study evaluated effects of teaching factors and extended day programs on mathematics learning, regarding both average achievement gains (quality dimension) and socioeconomic gaps in achievement gains (equity dimension). In the following, results of the empirical analyses are discussed.

The positive effect of *classroom management* on average achievement

Table 2
Intercept-and-slope-as-outcome models for effects extended day programs on mathematics final achievement (unstandardized regression coefficients with standard errors in brackets).

	Model 2a	Model 2b	Model 2c	Model 2d
Intercept	1.567 *	1.635 *	1.633 *	1.618 *
	(0.032)	(0.033)	(0.033)	(0.033)
Predictors student level				
SES (mean slope)	-0.055	0.039 *	0.039 *	0.038 *
	(0.054)	(0.011)	(0.011)	(0.011)
Mathematics prior achievement	0.539 *	0.539 *	0.537 *	0.542 *
	(0.026)	(0.026)	(0.027)	(0.027)
Migration background	-0.011	-0.024	-0.013	-0.015
	(0.061)	(0.063)	(0.064)	(0.063)
Gender	-0.227 *	-0.230 *	-0.232 *	-0.223 *
	(0.048)	(0.047)	(0.047)	(0.047)
Predictors class level: intercept				
Type of school (reference: half-day school)				
Half-day school with limited extended day program	0.135			
	(0.263)			
Open all-day school	0.053			
	(0.244)			
Partly bound all-day school	-0.152			
	(0.257)			
Fully bound all-day school	0.084			
	(0.243)			
Tutoring		0.091		
		(0.061)		
Homework supervision			-0.055 *	
			(0.018)	
Enrichment programs				-0.032^{+}
				(0.018)
Mean SES	0.045	0.029	0.045	0.047
	(0.041)	(0.041)	(0.041)	(0.038)
Predictors class level: slope				
Type of school (reference: half-day school)				
Half-day school with limited extended day program	0.108^{+}			
	(0.058)			
Open all-day school	0.107^{+}			
	(0.056)			
Partly bound all-day school	0.103^{+}			
	(0.057)			
Fully bound all-day school	0.045			
	(0.056)			
Tutoring		-0.046 *		
		(0.016)		
Homework supervision			0.025 *	
			(0.012)	
Enrichment programs				0.013 *
				(0.006)
Variance components				
Residual variance mathematics final achievement student level	0.616 *	0.617 *	0.620 *	0.621 *
	(0.027)	(0.028)	(0.027)	(0.028)
Residual variance mathematics final achievement class level	0.044^{+}	0.045 *	0.037 *	0.046 *
	(0.022)	(0.021)	(0.019)	(0.022)
Residual variance slope	0.001	0.001	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Covariance intercept slope	-0.003	-0.001	-0.001	-0.001
	(0.003)	(0.004)	(0.003)	(0.003)
Deviance test	18.789 *	66.966 *	105.789 *	100.301 *

Notes. $^{+}p < 0.10$, $*p < 0.05$.

gains is theoretical plausible and in line with previous research (Marzano et al., 2003). However, it was hypothesized that classroom management would reduce socioeconomic gaps in achievement gains, too. This was not confirmed (similar to Atlay et al., 2019b). An explanation might be that the analyzed variable reflects a fairly global measurement of classroom management. It might be that equalizing effects can only be detected if the measurement of classroom management focuses strongly on teachers' behavioral management strategies (e.g., setting clear rules, appropriate reactions to disruptive behavior).

The effects of *cognitive activation* on average achievement gains comprise both a linear and a quadratic effect. A similar pattern was found by Caro et al. (2016). This suggests that the optimal degree of cognitive activation is high, but not at the maximum. Further, cognitive activation does not affect socioeconomic achievement gaps meaning that higher-SES students and lower-SES students profit equally. This stands in contrast to the findings of Atlay et al. (2019b), who also examine secondary schools in Germany. The differing results might emerge because the current study focuses solely on academic track secondary schools while the study of Atlay et al. comprises several types of secondary schools (that is, both academic track and non-academic track secondary schools).

The analyses provide some indications of equalizing effects of *teachers' academic support*. While there is a tendency that tasks with academic support by the teacher come along with lower socioeconomic gaps in achievement gains, the same is not true for tasks without academic support by the teacher. This pattern of results is in line with the assumption that teachers' academic support reduces socioeconomic gaps in achievement gains.

As expected, forming *heterogenous groups* within the classroom reduces socioeconomic gaps in achievement gains. The same is not true for homogenous groups. This result is consistent with previous findings (Lou et al., 1996; Rohrbeck et al., 2003). Once again it is apparent that lower-SES students are in great need of stimulating interaction partners during learning.

Regarding the effects of extended day programs, several results are surprising. Concerning *types of schools*, there is a tendency that socioeconomic gaps in achievement gains are stronger in open and partly bound all-day schools than in half-day schools. Even if this effect is only marginally significant, it stands in stark contrast to political intentions. A plausible explanation might be that higher-SES students more often participate in extended day programs than lower-SES students. The expected equalizing effect of fully bound all-day schools is not confirmed, yet fully bound all-day schools do not increase socioeconomic gaps in achievement gains. Thus, within the group of all-day schools, fully bound all-day schools perform best in terms of the equity dimension. This is in line with the findings of Fischer et al. (2014).

Likewise surprising are the results on effects of *homework supervision*. A higher frequency of homework supervision is associated with lower average achievement gains. As the possibility of reverse causality can be excluded, reasons for adverse effects on learning need to be elaborated. It might be that the quality of school-based homework supervisions is low (e.g., large groups of students being supervised by staff with inadequate qualifications). In addition, doing homework in a supervised school setting might impair self-regulated learning, as compared to doing homework at home. Further questions are raised by the result that homework supervision is associated with larger socioeconomic gaps in academic achievement. This means that adverse effects of homework supervision are especially strong for lower-SES students. While it was theoretically expected that differences in parents' academic support play a role, this result rather suggests that differences in students' learning prerequisites are relevant. Specifically, higher-SES students might have the learning prerequisites necessary to adapt to school-based homework supervisions (e.g., focus on tasks, make use of school staff), but lower-SES might lack these learning prerequisites.

The results on *tutoring* confirm the hypothesis that tutoring reduces socioeconomic gaps in achievement gains, which is in line with previous findings on disadvantaged students (Fryer & Howard-Noveck, 2020). The equalizing effect is highly plausible given that tutoring addresses low-achieving students.

In the same vein, it is plausible that *enrichment programs* widen socioeconomic gaps in achievement gains as these programs address high-achieving students. However, there is also an unexpected tendency that enrichment programs reduce average achievement gains. It might be that highly frequent enrichment programs imply that schools allocate

considerable resources to a small group of students. This might come at the price of a lack of resources for the promotion of other students.

Finally, the results have some implications for the proposed mechanisms underlying the effects of teaching/school factors on socioeconomic achievement gaps (Section 3.1). It was proposed that effects of teaching/school factors on socioeconomic achievement gaps may be due either do *differentially effective learning opportunities* or to *different learning opportunities*. For most class level predictors that have significant effects on socioeconomic achievement gaps, it seems highly plausible that the mechanism of different learning opportunities is at work. For instance, tasks with academic support might affect socioeconomic achievement gaps because teachers treat low-SES students more supportively than high-SES students (Atlay et al., 2019a). In a similar vein, extended day programs might affect socioeconomic achievement gaps because low-SES students and high-SES students differ in participation rates (Linberg et al., 2015). From this, it can be hypothesized that different learning opportunities are a more powerful mechanism than differentially effective learning opportunities. Unfortunately, this assumption could not be tested empirically (see next section).

7.1. Limitations

Several limitations of the current study need to be acknowledged. First, it should be noted that the analyses refer to students in academic track secondary schools. This type of secondary school was chosen in order to have a socially heterogenous sample. On the downside, academic track secondary schools differ in several regards from other types of secondary schools in Germany. Therefore, it is not certain whether the obtained results are valid for other types of secondary schools, too. The same is true in terms of the transferability to primary schools.

A second limitation of the current study is the measurement of teaching factors/extended day programs. All variables stem either from teacher questionnaires or principal questionnaires. Clearly, data from student questionnaires and observational data is lacking. As all information on teaching factors/extended day programs is coming from school staff, only a narrow and one-sided perspective on the school learning environment is offered. It should also be noted that school staff self-reports are prone to bias (Lüdtke, Robitzsch, Trautwein, & Kunter, 2009). For instance, it might be that some teachers report a higher degree of academic support than they actually provide, either because they think it is expected of them (social-desirability bias) or because they want to protect their self-esteem (self-serving bias). In this case, an underestimation of the effects of academic support could occur.

A third limitation refers to the specific variables contained in the teacher and principal questionnaires. Several teaching factors/extended day programs are captured only in terms of quantitative aspects (e.g., time proportions, frequencies). However, qualitative aspects are considered equally important as quantitative aspects (Creemers & Kyriakides, 2008). For instance, the current study examines the time proportion for classroom management, but not the quality of teachers' classroom management strategies. Further, the support dimension is operationalized only in terms of teachers' academic support. This narrow conceptualization does not take into account other subdimensions of teacher support (e.g., social-emotional support) and support provided by students.

Finally, the study proposed mechanisms underlying the effects of teaching/school factors on socioeconomic achievement gaps. These mechanisms were used for the theory-driven identification of teaching/school factors that might affect socioeconomic achievement gaps. However, the study did not test these mechanisms empirically as respective analyses would require data on manifestations of the teaching/school factors at the student level (e.g., the degree of academic support by the teacher that individual students receive).

7.2. Implications for research, policy, and practice

The current study demonstrates the importance of considering both the quality and the equity dimension in evaluations of educational effectiveness. To date, many studies consider only average learning outcomes. However, this approach might overlook important dynamics of educational effectiveness, as certain school factors might be effective for some students but not for others (Creemers & Kyriakides, 2008). It is therefore recommended to standardly consider differences between students with various background characteristics. This not only enables a more sophisticated understanding of educational effectiveness, but also extends the knowledge base on interventions to reduce existing learning gaps.

This knowledge base is in many areas still insufficient. Regarding regular class, the roles that some teaching factors play for equity in education are not yet clarified. The reasons why effects of teaching factors on socioeconomic achievement gaps are found only in certain instances need to be investigated in more depth. In terms of classroom management, it should be examined whether measurements of teachers' behavioral management strategies are more consistently associated with socioeconomic achievement gaps than global measurements of classroom management. Regarding teacher support, it would be relevant to know which type of teacher support is most effective at reducing socioeconomic achievement gaps (e.g., academic vs. social-emotional support, proactive vs. reactive support). In a similar vein, the conditions under which cognitive activation increases socioeconomic achievement gaps need to be clarified (e.g., differences between school tracks). Regarding extended day programs, school-based homework supervision deserves more attention. While this type of program is widespread, empirical evidence on its effectiveness is lacking and the current study even gives indications of negative effects. All in all, more research on associations between teaching/school factors and gaps in learning outcomes is needed. To overcome limitations of the current study, it is recommended to use different data sources (school staff, students, observations) and to consider both quantitative and qualitative aspects of teaching/school factors. Student level variables on teaching/school factors seem especially valuable as they make it possible to examine the proposed mechanisms underlying the effects of teaching/school factors on socioeconomic achievement gaps. With this kind of data, interaction effects at the student level (differentially effective learning opportunities) as well as correlations between SES and learning opportunities (different learning opportunities) could be analyzed. It would even be possible to test whether within-school relationships between SES und learning opportunities are associated with socioeconomic achievement gaps (which could be called a slope-on-slope model).

The current study has implications for policy and practice, too. The results provide some evidence that certain school-based interventions are effective at reducing socioeconomic achievement gaps. The following recommendations can be given: 1) Teachers should provide a high degree of academic support in class. 2) Teachers should form groups of students with different achievement levels in the classroom. 3) Schools should offer tutoring in the afternoon hours.

Socioeconomic achievement gaps remain are major challenge for educational policy and practice. While the current study made a relevant contribution, there is still a long way to a deep understanding of how schools affect equity in education.

CRediT authorship contribution statement

Max Nachbauer: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization.

Acknowledgements

This paper uses data from the National Educational Panel Study (NEPS; see Blossfeld & Roßbach, 2019). The NEPS is carried out by the

Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.stueduc.2024.101367.

References

Atlay, C., Tieben, N., Fauth, B., & Hillmert, S. (2019a). The role of socioeconomic background and prior achievement for students' perception of teacher support. *British Journal of Sociology of Education, 40*(*7*), 970–991. https://doi.org/10.1080/01425692.2019.1642737

Atlay, C., Tieben, N., Hillmert, S., & Fauth, B. (2019b). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction, 63*. https://doi.org/10.1016/j.learninstruc.2019.05.008

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(*1*), 133–180. https://doi.org/10.3102/0002831209345157

Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, P. Schneider et al. (Ed.), PISA 2000 (pp. 323–407). Opladen.

Blossfeld, H.-P., & Roßbach, H.-G. (Eds.). (2019). Education as a lifelong process: The German National Educational Panel Study (NEPS). Springer.

Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. John Wiley & Sons,.

Brauns, H., Scherer, S., & Steinmann, S. (2003). The CASMIN educational classification in international comparative research. In J. H. P. Hoffmeyer-Zlotnik, & C. Wolf (Eds.), *Advances in cross-national comparison* (pp. 221–244). Springer.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328–375). Macmillan.

Caro, D. H., Lenkeit, J., & Kyriakides, L. (2016). Teaching strategies and differential effectiveness across learning contexts: Evidence from PISA 2012. *Studies in Educational Evaluation, 49*, 30–41. https://doi.org/10.1016/j.stueduc.2016.03.005

Charalambous, E., Kyriakides, L., & Creemers, B. P. M. (2018). Promoting quality and equity in socially disadvantaged schools: A group-randomisation study. *Studies in Educational Evaluation, 57*, 42–52. https://doi.org/10.1016/j.stueduc.2016.06.001

Cooper, H., Nye, B., & Lindsay, J. J. (2000). Homework in the home: How student, family, and parenting-style differences relate to the homework process. *Contemporary Educational Psychology, 25*(4), 464–487. https://doi.org/10.1006/ceps.1999.1036

Cosden, M., Morrison, G., Gutierrez, L., & Brown, M. (2004). The effects of homework programs and after-school activities on school success. *Theory into Practice, 43*(*3*), 220–226. https://doi.org/10.1207/s15430421tip4303_8

Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(*2*), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

European Commission/EACEA/Eurydice (2020). Equity in school education in Europe: Structures, policies and student performance. Eurydice report. Publications Office of the European Union.

Fauth, B., Atlay, C., Dumont, H., & Decristan, J. (2021). Does what you get depend on who you are with? Effects of student composition on teaching quality. *Learning and Instruction, 71*, Article 101355. https://doi.org/10.1016/j.learninstruc.2020.101355

Fischer, N., Theis, D., & Züchner, I. (2014). Narrowing the gap? The role of all-day schools in reducing educational inequality in Germany. *International Journal for Research on Extended Education, 2*(*1*), 79–96. https://doi.org/10.3224/ijree.v2i1.19535

Fryer, R., & Howard-Noveck, M. (2020). High-dosage tutoring and reading achievement: Evidence from New York City. *Journal of Labor Economics, 38*, 421–452. https://doi.org/10.3386/w23792

E. Garcia Inequalities at the Starting gate: Cognitive and noncognitive skills gaps between 2010–2011 Kindergarten classmates 2015.⟨https://files.eric.ed.gov/fulltext/ED560407.pdf⟩.

Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation, 57*, 16–30. https://doi.org/10.1016/j.stueduc.2016.09.004

Helmke, A. (1988). Leistungssteigerung und Ausgleich von Leistungsunterschieden in Schulklassen: Unvereinbare Ziele? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 20*, 45–76.

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (pp. 231–252). Erlbaum.

Karlen, Y., Merki, K. M., & Ramseier, E. (2014). The effect of individual differences in the development of metacognitive strategy knowledge. *Instructional Science, 42*(*5*), 777–794. https://doi.org/10.1007/s11251-014-9314-9

Kim, M. (2016). A meta-analysis of the effects of enrichment programs on gifted students. *Gifted Child Quarterly, 60*(*2*), 102–116. https://doi.org/10.1177/0016986216630607

König, J., Blömeke, S., Jentsch, A., Schlesinger, L., Felske, C., Musekamp, F., & Kaiser, G. (2021). The links between pedagogical competence, instructional quality, and mathematics achievement in the lower secondary classroom. *Educational Studies in Mathematics, 107*(*1*), 189–213. https://doi.org/10.1007/s10649-020-10021-0

Kyriakides, L., Charalambous, E., Creemers, B. P. M., & Dimosthenous, A. (2019a). Improving quality and equity in schools in socially disadvantaged areas. *Educational Research, 61*(*3*), 274–301. https://doi.org/10.1080/00131881.2019.1642121

Kyriakides, L., Creemers, B. P. M., & Charalambous, E. (2018). *Equity and quality dimensions in educational effectiveness*. Springer,. https://doi.org/10.1007/978-3-319-72066-1

Kyriakides, L., Creemers, B. P. M., & Charalambous, E. (2019b). Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity. *School Effectiveness and School Improvement, 30*(*3*), 286–308. https://doi.org/10.1080/09243453.2018.1511603

Linberg, T., Struck, O., & Bäumer, T. (2015). *Analysen zum Zusammenhang von ganztägiger Beschulung mit den Kompetenzen von Schülern in Deutschland*. Bertelsmann Stiftung,.

Linberg, T., Struck, O., & Bäumer, T. (2018). Vorzug Ganztagsschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft, 21*(6), 1205–1227. https://doi.org/10.1007/s11618-018-0830-2

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the pythagorean theorem. *Learning and Instruction, 19*(*6*), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Liu, J., Peng, P., Zhao, B., & Luo, L. (2022). Socioeconomic status and academic achievement in primary and secondary education: a meta-analytic review. *Educational Psychology Review, 34*(*4*), 2867–2896. https://doi.org/10.1007/s10648-022-09689-y

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta-analysis. *Review of Educational Research, 66*(*4*), 423–458. https://doi.org/10.3102/00346543066004423

Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology, 34*, 120–131. https://doi.org/10.1016/j.cedpsych.2008.12.001

Luyten, H. (2003). The Size of School Effects Compared to Teacher Effects: An Overview of the Research Literature. *School Effectiveness and School Improvement, 14*(*1*), 31–51. https://doi.org/10.1076/sesi.14.1.31.13865

Marzano, R.J., Marzano, J.S., & Pickering, D.J. (2003). Classroom management that works: Research-based strategies for every teacher. http://perino.pbworks.com/f/4%20CLM-Basic-Text(M).pdf.

Meyer, E., & Van Klaveren, C. (2013). The effectiveness of extended day programs: Evidence from a randomized field experiment in the Netherlands. *Economics of Education Review, 36*, 1–11. https://doi.org/10.1016/j.econedurev.2013.04.002

Miech, R., Essex, M. J., & Goldsmith, H. H. (2001). Socioeconomic status and the adjustment to school: The role of self-regulation during early childhood. *Sociology of Education, 74*(*2*), 102. https://doi.org/10.2307/2673165

Muthén, L.K. & Muthén, B. (1998–2018). Mplus user's guide. Los Angeles: Muthén & Muthén.

Nachbauer, M. (2023). Die Effekte von Schule auf Leistungsentwicklung und Leistungsunterschiede nach sozialer Herkunft. Eine Längsschnittstudie zu Ursachen von und Maßnahmen gegen Bildungsungleichheiten. Waxmann. https://doi.org/10.31244/9783830997320.

Nachbauer, M., & Kyriakides, L. (2020). A review and evaluation of approaches to measure equity in educational outcomes. *School Effectiveness and School Improvement, 31*(*2*), 306–331. https://doi.org/10.1080/09243453.2019.1672757

NEPS Network (2023). *National Educational Panel Study, scientific use file of starting cohort grade 5*. Leibniz Institute for Educational Trajectories (LIfBi), Bamberg. https://doi.org/10.5157/NEPS:SC3:12.1.0.

A. Nickow P. Oreopoulos V. Quan The impressive effects of tutoring on PreK-12 learning: A systematic review and meta-analysis of the experimental evidence 2020 doi: 10.3386/w27476.

Nilsen, T., Scherer, R., Gustafsson, J.-E., Teig, N., & Kaarstein, H. (2020). Teachers' role in enhancing equity: A multilevel structural equation modeling with mediated moderation. In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, Equality and Diversity in the Nordic Model of Education* (pp. 173–196). Springer.

Nuttall, D. L., Goldstein, H., Prosser, R., & Rasbash, J. (1989). Differential School Effectiveness. *International Journal of Educational Research, 13*(*7*), 769–776.

OECD. (2016). *Who uses childcare? Background brief on inequalities in the use of formal early childhood education and care (ECEC) among very young children*. OECD.

Opdenakker, M.-C., & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal, 27*(*4*), 407–432. https://doi.org/10.1080/01411920120071434

Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement, 19*(*1*), 21–49. https://doi.org/10.1080/09243450801936845

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM Mathematics Education, 50*(*3*), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Sage.

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual

reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63–72. https://doi.org/10.1016/j.learninstruc.2014.01.007

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240–257. https://doi.org/10.1037/0022-0663.95.2.240

Sammons, P. (1999). *School effectiveness. Coming of age in the twenty-first century*. Swets & Zeitlinger.

Scheerens, J. (2013). The use of theory in school effectiveness research revisited. *School Effectiveness and School Improvement, 24*(1), 1–38. https://doi.org/10.1080/09243453.2012.691100

Scheerens, J., & Bosker, R. J. (1997). *The Foundations of Educational Effectiveness*. Pergamon.

Steinmann, I., Strietholt, R., & Caro, D. (2019). Participation in extracurricular activities and student achievement: evidence from German all-day schools. *School Effectiveness and School Improvement, 30*(2), 155–176. https://doi.org/10.1080/09243453.2018.1540435

Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagsschulen. *Zeitschrift Für Erziehungswissenschaft, 18*(4), 737–761. https://doi.org/10.1007/s11618-015-0634-6

Tao, Y., Meng, Y., Gao, Z., & Yang, X. (2022). Perceived teacher support, student engagement, and academic achievement: a meta-analysis. *Educational Psychology, 42*(4), 401–420. https://doi.org/10.1080/01443410.2022.2033168

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. United Nations.

Zwier, D., Geven, S., & van de Werfhorst, H. G. (2021). Social inequality in shadow education: The role of high-stakes testing. *International Journal of Comparative Sociology, 61*(6), 412–440. https://doi.org/10.1177/0020715220984500